

University of Groningen

Reliability of ultrasonography in detecting shoulder disease in patients with rheumatoid arthritis

Bruyn, G. A. W.; Naredo, E.; Moeller, I.; Moragues, C.; Garrido, J.; de Bock, G. H.; d'Agostino, M-A; Filippucci, E.; Iagnocco, A.; Backhaus, M.

Published in:
Annals of the Rheumatic Diseases

DOI:
[10.1136/ard.2008.089243](https://doi.org/10.1136/ard.2008.089243)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2009

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Bruyn, G. A. W., Naredo, E., Moeller, I., Moragues, C., Garrido, J., de Bock, G. H., d'Agostino, M-A., Filippucci, E., Iagnocco, A., Backhaus, M., Swen, W. A. A., Balint, P., Pineda, C., Milutinovic, S., Kane, D., Kaeley, G., Narvaez, F. J., Wakefield, R. J., Narvaez, J. A., ... de Agustin, J. (2009). Reliability of ultrasonography in detecting shoulder disease in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, 68(3), 357-361. <https://doi.org/10.1136/ard.2008.089243>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Reliability of ultrasonography in detecting shoulder disease in patients with rheumatoid arthritis

G A W Bruyn, E Naredo, I Möller, C Moragues, J Garrido, G H de Bock, M-A d'Agostino, E Filippucci, A Iagnocco, M Backhaus, W A A Swen, P Balint, C Pineda, S Milutinovic, D Kane, G Kaeley, F J Narvaez, R J Wakefield, J A Narvaez, J de Augustin and W A Schmidt

Ann Rheum Dis 2009;68:357-361; originally published online 4 Apr 2008;
doi:10.1136/ard.2008.089243

Updated information and services can be found at:

<http://ard.bmj.com/cgi/content/full/68/3/357>

These include:

References

This article cites 23 articles, 10 of which can be accessed free at:
<http://ard.bmj.com/cgi/content/full/68/3/357#BIBL>

Rapid responses

You can respond to this article at:
<http://ard.bmj.com/cgi/eletter-submit/68/3/357>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

Topic collections

Articles on similar topics can be found in the following collections

[Immunology \(including allergy\)](#) (14050 articles)
[Radiology](#) (6335 articles)
[Connective tissue disease](#) (2469 articles)
[Degenerative joint disease](#) (2984 articles)
[Musculoskeletal syndromes](#) (4550 articles)
[Rheumatoid arthritis](#) (1686 articles)
[Clinical diagnostic tests](#) (8987 articles)
[Radiology \(diagnostics\)](#) (4930 articles)

Notes

To order reprints of this article go to:

<http://journals.bmj.com/cgi/reprintform>

To subscribe to *Annals of the Rheumatic Diseases* go to:

<http://journals.bmj.com/subscriptions/>

Reliability of ultrasonography in detecting shoulder disease in patients with rheumatoid arthritis

G A W Bruyn,¹ E Naredo,² I Möller,³ C Moragues,⁴ J Garrido,⁵ G H de Bock,⁶ M-A d'Agostino,⁷ E Filippucci,⁸ A Iagnocco,⁹ M Backhaus,¹⁰ W A A Swen,¹¹ P Balint,¹² C Pineda,¹³ S Milutinovic,¹⁴ D Kane,¹⁵ G Kaeley,¹⁶ F J Narvaez,⁴ R J Wakefield,¹⁷ J A Narvaez,¹⁸ J de Augustin,¹⁹ W A Schmidt²⁰

¹ Department of Rheumatology, Medisch Centrum Leeuwarden, Leeuwarden, The Netherlands;

² Department of Rheumatology, Hospital Severo Ochoa, Madrid, Spain; ³ Department of Rheumatology, Instituto Poal de Reumatologia, Barcelona, Spain;

⁴ Department of Rheumatology, Hospital de Bellvitge, Barcelona, Spain; ⁵ Department of

Methodology of Psychology, Universidad Autonoma de Madrid, Madrid, Spain;

⁶ Department of Epidemiology, University of Groningen, Groningen, The Netherlands;

⁷ Department of Rheumatology, Hopital Ambroise Pare, Paris, France; ⁸ Department of

Rheumatology, Università della Politecnica della Marche, Ancona, Italy; ⁹ Department of

Rheumatology, Università de la Sapienza, Rome, Italy; ¹⁰ Charite University Hospital, Berlin, Germany; ¹¹ Department of

Rheumatology, Medisch Centrum Alkmaar, Alkmaar, The Netherlands; ¹² Department of

Rheumatology, National Institute of Rheumatology and Rehabilitation, Budapest, Hungary; ¹³ National Institute of

Rehabilitation, Mexico City, Mexico; ¹⁴ Institute for Rheumatology and

Rehabilitation, Belgrade, Serbia; ¹⁵ Department of Rheumatology, Adelaide and Meath Hospital, Dublin, Ireland; ¹⁶ ROAD Clinic, Lakewood, Washington, USA;

¹⁷ Department of Rheumatology, Chapel Allerton Hospital, Leeds, UK; ¹⁸ Department of Radiology, Hospital de Bellvitge, Barcelona, Spain; ¹⁹ Department of

Rheumatology, Hospital Val d'Hebron, Barcelona, Spain;

²⁰ Department of Rheumatology, Hospital Berlin-Buch, Berlin, Germany

Correspondence to: G A W Bruyn, Department of Rheumatology, Medisch Centrum Leeuwarden, 8934 AD Leeuwarden, Leeuwarden, The Netherlands; gawbruyn@wxss.nl

Accepted 30 March 2008
Published Online First 4 April 2008

ABSTRACT

Objective: To assess the intra and interobserver reproducibility of musculoskeletal ultrasonography (US) among rheumatologists in detecting destructive and inflammatory shoulder abnormalities in patients with rheumatoid arthritis (RA) and to determine the overall agreement between US and MRI.

Methods: A total of 14 observers examined 5 patients in 2 rounds independently and blindly of each other. US results were compared with MRI. Overall agreement of all findings, of positive findings on MRI, as well as intra and interobserver reliabilities, were calculated.

Results: Overall agreement between US and MRI was seen in 79% with regard to humeral head erosions (HHE), in 64% with regard to posterior recess synovitis (PRS), in 31% with regard to axillary recess synovitis (ARS), in 64% with regard to bursitis, in 50% with regard to biceps tenosynovitis (BT), and in 84% for complete cuff tear (CCT). Intraobserver and interobserver κ was 0.69 and 0.43 for HHE, 0.29 and 0.49 for PRS, 0.57 and 1.00 for ARS, -0.17 and 0.51 for bursitis, 0.17 and 0.46 for BT and 0.52 and 0.6 for CCT, respectively. The intraobserver and interobserver κ for power Doppler (PD) was 0.90 and 0.70 for glenohumeral signals and 0.60 and 0.51 for bursal signals, respectively.

Conclusions: US is a reliable imaging technique for most shoulder pathology in RA especially with regard to PD. Standardisation of scanning technique and definitions of particular lesions may further enhance the reliability of US investigation of the shoulder.

Shoulder involvement is a challenging issue in rheumatoid arthritis (RA), bringing about a deleterious impact on the quality of life in many of those affected by the disease.¹⁻⁵ In RA, ongoing synovial inflammation may lead not only to erosive shoulder disease, but also to rotator cuff rupture. Taking the severity of morbidity and the serious complications of shoulder pathology for patients with RA into account, detecting synovitis at an early stage is a key issue for prevention of irreversible damage.

Ultrasonography (US) is an imaging modality that is now widely accepted in rheumatology research and clinical practice to visualise joints and soft tissues in patients with various rheumatic diseases. US is able to not only image the damage to cartilage and bone, but also to identify tendon pathology and synovial inflammation. Patients are likely to have shoulder tendon disease if US is abnormal.⁶⁻⁸ Despite the increasing use of US, the

technique is regarded as examiner-dependent. Furthermore, notwithstanding increasing data on the reliability of US in the evaluation of small joints of the hand and the feet, there is a clear paucity of studies regarding reliability of US for other joints.⁹⁻¹²

In the light of these limitations, we undertook a first step in investigating these issues for patients with established RA and shoulder disease. We addressed the agreement between US and MRI, as well as the intra and inter-reader variability among rheumatologists with experience in musculoskeletal US.

PATIENTS AND METHODS

Patient selection

Five patients with symptomatic shoulder disease and RA were selected from the Barcelona University Rheumatology Unit. There were two men and three women, with a median age of 64.8 (55 to 76) years and median disease duration of 6.6 (1 to 10) years. All had established RA according to the American College of Rheumatology (ACR) 1987 criteria for RA. All patients were investigated twice (ie, the procedure was repeated during the afternoon session), with rearrangement of the patients in a different order and on a different location.

Observers

Observers consisted of a group of 14 rheumatologists from 9 countries with variable expertise (median experience 10 years, range 3 to 16 years) in musculoskeletal US. All were members of the Outcome Measures in Rheumatology (OMERACT) US group. The observers met for 1 day to perform the investigation. The sonographers performed the US investigation independently from each other. The observers were blinded to the clinical details and MRI results. All investigators met for a brief training session before the exercise, to review the scoring method and for initial training of observers not familiar with some aspects of the scoring system or the US machine. A statistician was on hand to receive the filled score sheets. The score sheets from the morning session were sealed in envelopes until the second session was concluded.

US imaging

All scans were performed using a Siemens Acuson Antares (Siemens, Erlangen, Germany) machine

Extended report

with a 7.5 to 15 MHz linear array transducer. The shoulder scoring system assessed elements of inflammation, as well as structural tendinous and bony damage. Rotator cuff tendons were investigated for the presence of total or partial tears in a longitudinal and a transverse plane on static and dynamic positions. The synovial structures of the shoulder including subacromial-subdeltoid bursa, sheath of the long biceps tendon and the axillary and posterior recess of the glenohumeral joint, were examined for the presence of effusions and synovial hypertrophy. The humeral head was examined for the presence of erosions. Power Doppler assessment of selected synovial sites including biceps sheath, subacromial-subdeltoid bursa and axillary and posterior recesses was carried out with settings standardised to a pulse repetition frequency of 400 to 500 Hz and low wall filters. The power Doppler gain was adjusted to a level just below the disappearance of colour signs under the bony cortex as recommended by Rubin *et al.*¹³ OMERACT definitions for joint effusion, synovial hypertrophy, tenosynovitis and bone erosions were adhered to.¹⁴ The following definitions for the classification of ultrasonographic findings were used: cortical irregularities >2 mm were considered as erosions; a hypoechoic area of at least 3 mm around the long head of the biceps tendon was considered a tenosynovitis of the long biceps tendon, bursal thickness >3 mm or effusion as effusion/synovial hypertrophy of the subacromial-subdeltoid bursa, >3 mm effusion/synovial hypertrophy at the posterior recess superior to the glenoid labrum as synovitis, >3 mm effusion/synovial hypertrophy at the axillary recess as synovitis. No ultrasonographic distinction was made between effusions and synovial hypertrophy and these abnormalities were taken together for the analyses.

MRI

Assessment of the affected shoulder by MRI took place within 5 working days prior to the US investigation in all patients. MR imaging was performed with a 1.5-T unit (Signa Excite, General Electric, Milwaukee, Wisconsin, USA) using a flexible wrap-around coil. The following sequences were used: T1-weighted spin-echo sequence (repetition time (TR) of 500 ms, echo time (TE) of 13.6 ms, slice thickness (SL) 4 mm and fields of view (FOV) of 140–160 mm in an axial, transverse and oblique coronal slice orientation parallel to the course of the tendon of the supraspinatus; T2-weighted fat suppressed images in a coronal and a sagittal plane with a TR of 3300 ms and a TE of 71 ms.

After selection of a suitable slice on which abnormal changes were visualised, a dynamic contrast-enhanced study gadolinium diethylenetriaminepentaacetic acid (GD-DTPA; 0.1 mmol/kg of body weight) was performed using an axial T1-weighted (TR of 500 ms, TE 13.6 ms and FOV 140–160 mm), axial contrast-enhanced fat suppressed T1-weighted and coronal oblique contrast-enhanced fat suppressed T1-weighted sequences were performed.

The MRI scans were evaluated by two radiologists who were in consensus and had no knowledge of the results of the ultrasonography. The MRI scans were analysed for the presence or absence of the same structures that were visualised by ultrasonography (table 1). The MRI criterion for effusion was an intra-articular or intrabursal area with a high signal on T2-weighted sequences without contrast enhancement on fat-suppressed T1-weighted sequences. The criterion for synovitis was enhancing material seen on the fat-suppressed T1-weighted sequences.¹⁵

Table 1 Blank copy of score sheet used for ultrasound (US) assessment of shoulder pathology in patients with rheumatoid arthritis (RA)

| Name | Date | Round I/round II |
|--|--------|------------------|
| Patient name | | |
| Pathological finding on US | | MRI confirmation |
| Biceps tenosynovitis (3 mm) | Yes/no | Yes/no |
| Tenosynovial power Doppler signal: | | |
| Synovitis: | | |
| Axillary recess (3 mm) | Yes/no | Yes/no |
| Posterior recess (3 mm) | Yes/no | Yes/no |
| Synovial power Doppler signal: | | |
| Axillary recess | Yes/no | Yes/no |
| Posterior recess | Yes/no | Yes/no |
| Bursitis: subdeltoid-subscapular bursa | Yes/no | Yes/no |
| Intrabursal power Doppler signal: | | |
| Partial cuff tear | | |
| Subscapularis | Yes/no | Yes/no |
| Supraspinatus | Yes/no | Yes/no |
| infrapinatus tendon | Yes/no | Yes/no |
| Complete tear: | | |
| Subscapularis | Yes/no | Yes/no |
| Supraspinatus | Yes/no | Yes/no |
| Infrapinatus tendon | Yes/no | Yes/no |
| Humeral head erosions | Yes/no | Yes/no |

Statistical analysis

Overall agreement between US and MRI was calculated for each observer. Averaged overall agreement and kappa index (κ) are shown. Since Cohen κ is artificially low in cases of high or low prevalence, we have used κ adjusted by prevalence and bias instead of κ standard.^{16 17} Furthermore, the mean positive and negative percentages of agreement were calculated.

Intraobserver reliability is presented as overall agreement between the first and second round for each scan and κ adjusted by prevalence and bias. Agreement indexes were interpreted as follows: 0.81–1.00 excellent agreement; 0.61–0.80, good agreement; 0.41–0.60, moderate agreement; 0.21–0.40, fair agreement; 0.00–0.20, slight agreement and <0.00, poor agreement.

Interobserver reliability was studied by calculating the generalised Scott π , also known as majority κ . The Scott π is the accepted standard for interobserver reliability for nominal data in communication studies, when there are more than two observers. Comparable to κ , π discounts the level of “observed agreement” by the level of “expected agreement” due to chance in the following way: $\pi = (\text{observed agreement} - \text{expected agreement}) / (1 - \text{expected agreement})$. Because there were 14 observers, agreement was declared (observed agreement) when at least 10 out of 14 observers assigned the same score to a given US scan. The probability of having at least 10 equal scores by chance was the expected agreement used to compute majority κ .

RESULTS

Validation

MRI investigation reported biceps tenosynovitis of the long biceps tendon, glenohumeral synovitis, subacromial-subdeltoid bursitis and humeral head erosions in four out of five patients with RA. Rotator cuff tears, either partial or complete, were seen on MRI in fewer patients. No total cuff tear of the infrapinatus tendon was found with MRI (table 2).

Table 2 Ultrasonography (US) definitions for various abnormalities

| Abnormality | Definition |
|----------------------|---|
| Joint effusion | Abnormal hypoechoic or anechoic intraarticular material that is displaceable and compressible |
| Synovial hypertrophy | Abnormal hypoechoic intraarticular tissue that is non-displaceable and poorly compressible, and that may exhibit Doppler signal |
| Tenosynovitis | Hypoechoic or anechoic thickened tissue within the tendon sheath, which is seen in two perpendicular planes |
| Erosion | An intraarticular discontinuity of the bone surface that is visible in two perpendicular planes |
| Bursitis | Abnormal hypoechoic or anechoic intrabursal material that is displaceable and compressible |
| Partial tendon tear | Partial-thickness interruption of tendon fibers with or without hypoechoic material filling the defect |
| Complete tendon tear | Full-thickness interruption of the tendon fibers with or without hypoechoic material filling the defect |

Table 3 lists the mean overall agreement (for either the presence or absence of pathological findings, ie, the accuracy) between the US observations and the MRI findings, as well as the positive agreement (ie, sensitivity, when the pathological finding is present on the MRI) and the negative agreement (ie, specificity, when the pathological finding was not present on the MRI). Table 2 demonstrates that good agreement between US and MRI was found for the presence/absence of erosions of the humeral head. Presence or absence of glenohumeral synovitis was found with moderate to good agreement. Regarding posterior recess synovitis, the mean agreement between MRI and US assessment was much better (64%) than for axillary recess synovitis (31%). Excellent agreement was found between US and MRI for the presence or absence of a complete rotator cuff tear, whereas with regard to partial cuff tears, US demonstrated a lower sensitivity.

Reproducibility

Table 4 lists the mean intraobserver agreement and the corresponding mean κ values. The mean overall agreements for intraobserver reproducibility ranged from moderate to excellent. Excellent intraobserver agreement was observed for humeral head erosions, complete and partial tears of subscapularis and infraspinatus tendon and power Doppler signals regarding the glenohumeral joint and bursa. The mean κ value

for intraobserver reproducibility for humeral head erosions was good (0.69). The mean κ values for synovitis were moderate (0.57) and fair (0.29) for axillary recess synovitis and posterior recess, respectively, whereas the κ value for bursitis was poor (−0.17). The mean κ values for synovial power Doppler flow in the joint recesses were excellent, whereas that for the power Doppler flow in the bursa was good. A poor κ value was found for long biceps tendon tenosynovitis. According to κ values, the intraobserver reproducibility for partial cuff tear ranged from slight to moderate, for total cuff tear from poor to moderate.

Table 5 lists the majority κ values for interobserver agreement. Increasing κ values signify better agreement between the 14 observers. The mean interobserver κ value for bony erosions was moderate (0.43). The κ value for tenosynovitis of the long biceps tendon was moderate (0.46), and a poor κ value for the power Doppler signal within the tendon sheath was found. Mean κ values for glenohumeral joint synovitis ranged from moderate (0.49) to excellent (1.0), with excellent interobserver mean majority κ values for the presence of power Doppler signal for either the axillary joint recess or the posterior recess. Interobserver κ for bursitis was moderate (0.51), with a negative mean κ value for PD signal within the bursa. The κ for partial cuff tear scored a good agreement, as well as the interobserver agreement for complete cuff tear.

Table 3 Agreement of ultrasonography (US) with MRI in five patients with rheumatoid arthritis

| Pathological finding | No. of positive findings on MRI in five patients | Overall agreement between US and MRI in first round (mean, %) | κ Adjusted by prevalence and bias (mean) | Positive agreement (mean, %) | Negative agreement (mean, %) |
|--|--|---|---|------------------------------|------------------------------|
| Biceps tenosynovitis | 4 | 50 | 0.086 | 63 | 15 |
| Synovitis: | | | | | |
| Axillary recess | 4 | 31 | −0.314 | 26 | 26 |
| Posterior recess | 4 | 64 | 0.371 | 71 | 49 |
| Bursitis of subdeltoid-subscapular bursa | 4 | 64 | 0.286 | 68 | 55 |
| Partial cuff tear: | | | | | |
| Subscapularis | 2 | 51 | 0.029 | 11 | 63 |
| Supraspinatus | 3 | 67 | 0.314 | 65 | 65 |
| Infraspinatus tendon | 3 | 46 | −0.057 | 41 | 48 |
| Complete tear: | | | | | |
| Subscapularis | 1 | 80 | 0.629 | 24 | 85 |
| Supraspinatus | 2 | 81 | 0.629 | 64 | 86 |
| Infraspinatus tendon | 0 | 91 | 0.829 | 0 | 95 |
| Humeral head erosions | 4 | 79 | 0.543 | 85 | 27 |

Extended report

Table 4 Intraobserver agreement of ultrasonography (US) of shoulder joint in rheumatoid arthritis (RA)

| Pathological finding | Pairs of observations | Overall agreement (mean, %)* | κ Adjusted by prevalence and bias (mean)* |
|--|-----------------------|------------------------------|--|
| Biceps tenosynovitis | 70 | 59 | 0.17 |
| Power Doppler tenosynovitis present | 70 | 51 | 0.03 |
| Synovitis: | | | |
| Axillary recess | 70 | 79 | 0.57 |
| Posterior recess | 70 | 64 | 0.29 |
| Synovial power Doppler signal: | | | |
| Axillary recess | 70 | 94 | 0.89 |
| Posterior recess | 60 | 96 | 0.91 |
| Bursitis: subdeltoid-subscapular bursa | 70 | 41 | −0.17 |
| Power Doppler bursa present | 70 | 80 | 0.60 |
| Partial cuff tear: | | | |
| Subscapularis | 70 | 80 | 0.60 |
| Supraspinatus | 70 | 57 | 0.14 |
| Infraspinatus tendon | 70 | 84 | 0.69 |
| Complete tear: | | | |
| Subscapularis | 70 | 81 | 0.63 |
| Supraspinatus | 70 | 64 | 0.29 |
| Infraspinatus tendon | 70 | 81 | 0.63 |
| Humeral head erosions | 70 | 84 | 0.69 |

*Mean of 14 intraobserver values, 1 for each observer.

DISCUSSION

This is the first study undertaken to date that focus on validation of ultrasonography-detected shoulder abnormalities in patients with established RA. We used MRI as the reference imaging technique. The results show that US can reliably assess evidence of joint destruction (ie, rotator cuff tears and erosions of the humeral head). In addition, our study shows that US can reliably assess signs of joint inflammation (eg, synovitis of the posterior recess and Doppler signals). This is also the first study to examine the US reproducibility of erosions and synovitis in the shoulder among rheumatologists-sonographers; however, the results indicate that for a limited number of shoulder changes, US still is an imaging technique with a wide spectrum of intra and interobserver variability.

Various studies on patients with shoulder disease in RA have demonstrated that US is comparable to MRI in being more sensitive than radiography in detecting bone erosion.^{18–21} One study showed that erosions were reported by US in 30 patients and by MRI in 39 patients, vs 26 by conventional radiography.^{18–19} The studies of Scheel *et al*²⁰ and Naredo *et al*²¹ demonstrated that an excellent agreement between US and MRI existed for humeral head erosions (84.5 and 100%, respectively), with a similar or lower detection rate of synovitis or effusion of the shoulder (88.5 and 50%, respectively). In our study, although radiography was not used as a comparator, we confirmed the high agreement level between US and MRI for detecting erosions of the humeral head. As to humeral head erosions detected by US, our mean κ values showed a good intraobserver and a moderate interobserver reliability.

We considered partial or complete cuff tear as a second outcome parameter of chronic damage of the shoulder in RA. We found an excellent agreement between the US and the MRI for complete cuff tear and, as expected, lesser agreement for partial cuff tear. The intra and interobserver agreements for partial and complete cuff tears varied from poor to excellent. Again, these results are consistent with earlier studies, ranging from 60% to 94%.^{6–7 22 23}

Furthermore, we compared the agreement of US and MRI in detecting inflammatory changes (ie, synovitis, bursitis and long

biceps tendon tenosynovitis). The detection of early inflammation of a joint is of key importance for the optimal management of patients with RA. Since the visualisation of synovial hypertrophy of the shoulder joint is extremely difficult with US due to its deep anatomic site, we did not actually diagnose synovitis by the detection of synovial hypertrophy, but by the presence of effusion and by the presence of power Doppler flow. As most synovitis—but not all—is accompanied by effusion, this probably introduces a small methodological error. All inflammatory conditions were found more frequently by MRI than by US. Our study was able to detect synovitis of the posterior process in over 60% of patients, indicating a high

Table 5 Interobserver agreement for ultrasonography (US) of the shoulder in rheumatoid arthritis (RA)

| Pathological finding | Observed majority agreement* (mean†, %) | Majority κ * |
|--|---|---------------------|
| Biceps tenosynovitis | 60.0 | 0.455 |
| Tenosynovitis power Doppler present | 40.0 | 0.183 |
| Synovitis: | | |
| Axillary recess | 100.0 | 1.000 |
| Posterior recess | 60.0 | 0.487 |
| Synovial power Doppler signal: | | |
| Axillary recess | 100.0 | 1.000 |
| Posterior recess | 100.0 | 1.000 |
| Bursitis: subdeltoid-subscapular bursa | 60.0 | 0.508 |
| Power Doppler bursa present | 80.0 | −1.621 |
| Partial cuff tear: | | |
| Subscapularis | 100.0 | 1.000 |
| Supraspinatus | 60.0 | 0.489 |
| Infraspinatus tendon | 100.0 | 1.000 |
| Complete tear: | | |
| Subscapularis | 100.0 | 1.000 |
| Supraspinatus | 60.0 | −0.238 |
| Infraspinatus tendon | 100.0 | 1.000 |
| Humeral head erosions | 80.0 | 0.434 |

*Agreement is declared if there are at least 10 equal scores out of 14 observers;

†mean of five interobserver values, one for each patient.

sensitivity, whereas the agreement for synovitis of the axillary recess was much lower. If the conclusion is justified that the posterior recess is more sensitive to synovitis, it would significantly shorten the US examination time.²⁴ With regard to bursitis, there was a high agreement between US and MRI, but a negative mean interobserver κ value. Moreover, the mean interobserver κ value for biceps tenosynovitis was poor. The poor and negative mean κ value mainly have technical reasons (ie, due to the high prevalence of the abnormality on the MRI the chance is relatively high that the US examination will effectively find the abnormality). Because the adjusted κ corrects for this high chance, the κ value will become negative when the observed US findings are less than the expected US findings.

These findings may lead to the conclusion that in cases of RA, signs of shoulder synovitis should be looked for at the posterior recess and not in the axillary recess. Furthermore, the intra and interobserver agreement for the power Doppler signal in the posterior and the axillary recess of the shoulder were excellent, suggesting that power Doppler signal may be used in multiple center studies as a parameter for active shoulder synovitis.

The lowest sensitivities of US were found for assessing axillary recess synovitis, long biceps tenosynovitis or partial cuff tears. It cannot be precluded that MRI overdiagnosed these findings (eg, some cases of tendonitis might have been interpreted as partial cuff tears). Although US is a dynamic investigation, thus improving its sensitivity to detect for example small quantities of fluid or small cuff tears, it was noted that not all observers fully used the dynamic approach. Moreover, some investigators were not familiar with the equipment and the scanner settings. The level of experience was also different for each sonographer, but we did not examine whether there was a correlation between the investigator's experience and his US performance. Perhaps a 10 min investigation, chosen because this amount of time reflected a busy clinical practice, was for some investigators too short to perform a thorough examination.

A limitation of our study is the small number of patients. Since there were only 5 available machines and 14 sonographers, the experiment was set up in such a way that as much information as possible would be obtained in a single working day. A larger sample including healthy persons and patients with various degrees of shoulder disease could yield more accurate information regarding the index of agreement between observers and test-retest reliability, but this would require a longer experiment.

In summary, this study shows that US is a reliable method for detection of erosions and complete cuff tear, and also reliable in detecting synovitis of the posterior recess of the glenohumeral joint and subdeltoid-subacromial bursitis. The 14 ultrasonographers/rheumatologists were able to detect these changes with a moderate to good interobserver reproducibility and similar intraobserver reproducibility. More studies are warranted focusing on improvement of US diagnosing of particularly biceps tenosynovitis, partial cuff tears and axillary synovitis.

Acknowledgements: The authors would like to thank Roche Pharmaceuticals Espana for funding the study. The authors would also like to thank Siemens for providing five US machines.

Funding: Roche Pharmaceuticals Espana funded this study.

Competing interests: None declared.

Ethics approval: Ethics approval was obtained.

REFERENCES

1. Tanaka E, Saito A, Kamitsuiji S, Yamahara T, Nakajima A, Tanijuchi A, *et al*. Impact of shoulder, elbow, and knee involvement on assessment of rheumatoid arthritis using the American College of Rheumatology Core Data Set. *Arthritis Rheum* 2005;**53**:864–71.
2. Petersson CJ. Painful shoulders in patients with rheumatoid arthritis. Prevalence, clinical and radiological features. *Scand J Rheumatol* 1986;**3**:437–51.
3. Alasareela E, Suramo I, Tervonen O, Lähde S, Takalo R, Hakala M. Evaluation of humeral head erosions in rheumatoid arthritis: a comparison of ultrasonography, magnetic resonance imaging, computed tomography and plain radiography. *Br J Rheum* 1998;**37**:1152–6.
4. Kuper HH, van Leeuwen MA, van Riel PL, Prevoo ML, Houtman PM, Lolkema WF, *et al*. Radiographic damage in large joints in early rheumatoid arthritis: relationship with radiographic damage in hands and feet, disease activity, and physical disability. *Br J Rheumatol* 1997;**36**:855–60.
5. Drossaers-Bakker KW, Kroon HM, Zwinderman AH, Breedveld FC, Hazes JMW. Radiographic damage of large joints in long-term rheumatoid arthritis and its relation to function. *Rheumatology* 2000;**39**:998–1003.
6. Bretzke CA, Crass JR, Craig EV, Feinberg SB. Ultrasonography of the rotator cuff. Normal and pathologic anatomy. *Invest Radiol* 1985;**20**:311–5.
7. Mack LA, Masten FA, Kilcynne RF, Davies PK, Sickler ME. US evaluation of the rotator cuff. *Radiology* 1985;**157**:205–9.
8. Wiener SN, Seitz WH. Sonography of the shoulder in patients with tears of the rotator cuff: accuracy and value for selecting surgical options. *Am J Roentgenol* 1993;**160**:103–7.
9. Balint PV, Sturrock RD. Intraobserver repeatability and interobserver reproducibility in musculoskeletal ultrasound imaging measurements. *Clin Exp Rheumatol* 2001;**19**:89–92.
10. Szkudlarek M, Court-Payen M, Jacobsen S, Klarlund M, Thomsen HS, Ostergaard M. Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. *Arthritis Rheum* 2003;**48**:955–62.
11. Quistgaard E, Torp-Pedersen S, Christensen R, Bliddal H. Reproducibility and inter-reader agreement of a scoring system for ultrasound evaluation of hip osteoarthritis. *Ann Rheum Dis* 2006;**65**:1613–9.
12. Karim Z, Wakefield RJ, Quinn M, Conaghan PG, Brown AK, Veale DJ, *et al*. Validation and reproducibility of ultrasonography in the detection of synovitis of the knee. A comparison with arthroscopy and clinical examination. *Arthritis Rheum* 2004;**50**:387–94.
13. Rubin JM, Adler RS, Fowlkes JB, Spratt S, Pallister JE, Chen JF, *et al*. Fractional moving blood volume: estimation with power Doppler US. *Radiology* 1995;**197**:183–90.
14. Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, D'Agostinho M, *et al*. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *J Rheumatol* 2005;**32**:2485–7.
15. Wamser G, Bohndorf K, Vollert K, Bucklein W, Schalm J. Power Doppler sonography with and without echo-enhancing contrast agent and contrast-enhanced MRI for the evaluation of rheumatoid arthritis of the shoulder joint: differentiation between synovitis and joint effusion. *Skeletal Radiol* 2003;**32**:351–9.
16. Byrt T, Bishop J, Carlin B. Bias, prevalence and kappa. *J Clin Epidemiol* 1993;**46**:423–9.
17. Craig TG. Generalization of Scott's Index of intercoder agreement. *Pub Opin Quarterly* 1981;**45**:260–4.
18. Hermann K-G, Backhaus M, Schneider U, Labs K, Loreck D, Zuhlsdorf S, *et al*. Rheumatoid arthritis of the shoulder joint. Comparison of conventional radiography, ultrasound, and dynamic contrast-enhanced magnetic resonance imaging. *Arthritis Rheum* 2003;**48**:3338–49.
19. Bruyn GA, Iagnocco A, Naredo E, Wakefield RJ, Schmidt WA. Ultrasonography of the shoulder in patients with rheumatoid arthritis: comment on the article by Hermann *et al*. *Arthritis Rheum* 2004;**50**:3054.
20. Scheel AK, Schmidt WA, Hermann KG, Bruyn GAW, D'Agostino MA, Grassi W, *et al*. Interobserver reliability of rheumatologists performing musculoskeletal ultrasonography: results from a EULAR train the trainers course. *Ann Rheum Dis* 2005;**64**:1043–9.
21. Naredo E, Moller I, Moragues C, De Agustin JJ, Scheel AK, Grassi W, *et al*. Interobserver reliability in musculoskeletal ultrasonography: results from a teach the teacher rheumatologist course. *Ann Rheum Dis* 2006;**65**:14–9.
22. Burk DL, Karasick D, Kurtz AB, Mitchell DG, Rifkin MD, Miller CL, *et al*. Rotator cuff tears: prospective comparison of MR imaging with arthrography, sonography, and surgery. *Am J Roentgenol* 1989;**153**:87–92.
23. Swen WAA, Jacobs JWG, Algra PR, Manoliu RA, Rijkman J, Lems WJ, *et al*. Sonography and magnetic resonance imaging equivalent for the assessment of full-thickness rotator cuff tears. *Arthritis Rheum* 1999;**42**:2231–8.
24. Schmidt WA, Schicke B, Krause A. Which ultrasound scan is the best to detect glenohumeral joint effusions? [in German] *Ultraschall Med* 2008. In press.